

[DASHBOARD](#) / [I MIEI CORSI](#) / [APPELLI DI CLAUDIO SARTORI](#) / [SEZIONI](#) / [MACHINE LEARNING](#) / [MACHINE LEARNING THEORY](#)

<b>Iniziato</b>	Monday, 7 February 2022, 14:50
<b>Stato</b>	Completato
<b>Terminato</b>	Monday, 7 February 2022, 15:08
<b>Tempo impiegato</b>	18 min. 21 secondi
<b>Punteggio</b>	15,00/15,00
<b>Valutazione</b>	<b>30,00</b> su un massimo di 30,00 ( <b>100%</b> )

Domanda **1**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

In data preprocessing, which of the following **is not** an objective of the *aggregation* of attributes

Scegli un'alternativa:

- ☒ a. Obtain a more detailed description of data
- ☐ b. Reduce the variability of data
- ☐ c. Reduce the number of attributes or objects
- ☐ d. Obtain a less detailed scale

Domanda **2**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Given the two binary vectors below, which is their similarity according to the Jaccard Coefficient?

**abcdefghij**

1000101101

1011101010

**Scegli un'alternativa:**☐ a. 0.1☒ b. 0.375☐ c. 0.5☐ d. 0.2Domanda **3**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Which of the following statements is *true*?

**Scegli una o più alternative:**☒ a. The data which are similar to the majority are never noise☐ b. The noise can generate outliers☐ c. Outliers can be due to noise☐ d. The noise always generate outliers

Domanda **4**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Given the definitions below:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

which of the formulas below computes the accuracy of a binary classifier?

**Scegli un'alternativa:**

- ☒ a.  $TP / (TP + FP)$
- ☐ b.  $(TP + TN) / (TP + FP + TN + FN)$
- ☐ c.  $TN / (TN + FP)$
- ☐ d.  $TP / (TP + FN)$

Domanda **5**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

What is the *cross validation*

**Scegli un'alternativa:**

- ☒ a. A technique to obtain a good estimation of the performance of a classifier with the training set
- ☒ b. A technique to obtain a good estimation of the performance of a classifier when it will be used with data different from the training set
- ☐ c. A technique to improve the speed of a classifier
- ☐ d. A technique to improve the quality of a classifier

Domanda **6**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

In a decision tree, the number of objects in a node...

**Scegli un'alternativa:**

- ☒ a. ...is smaller than or equal to the number of objects in its ancestor
- ☐ b. ...is smaller than the number of objects in its ancestor
- ☐ c. ...is bigger than the number of objects in its ancestor
- ☐ d. ...is not related to the number of objects in its ancestor

Domanda **7**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

When training a neural network, what is the *learning rate*?

- ☒ a. A multiplying factor of the correction to be applied to the connection weights
- ☐ b. The slope of the activation function in a specific node
- ☐ c. The ratio between the size of the hidden layer and the input layer of the network
- ☐ d. The speed of convergence to a stable solution during the learning process

Domanda 8

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Which of the statements below about *Hierarchical Agglomerative Clustering* is true?

- ☒ a. Requires the definition of *distance between sets of objects*
- ☐ b. Is based on a well founded statistical model
- ☐ c. Requires the definition of *Inertia* of clusters
- ☐ d. Is very efficient, also with large datasets

Domanda 9

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Which of the following *is not* a strength point of *Dbscan* with respect to *K-means*

**Scegli un'alternativa:**

- ☒ a. The *effectiveness*, even in presence of *noise*
- ☐ b. The efficiency even in large datasets
- ☐ c. The *robustness* with respect to outliers
- ☐ d. The *effectiveness* even if there are clusters with non-convex shape

Domanda **10**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Which of the statements below is true? (One or more)

**Scegli una o più alternative:**

- ☒ a. DBSCAN always stops to a configuration which gives the optimal number of clusters
- ☒ b. Increasing the radius of the neighbourhood can decrease the number of noise points
- ☒ c. DBSCAN can give good performance when clusters have concavities
- ☒ d. Sometimes DBSCAN stops to a configuration which does not include any cluster

Domanda **11**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Which of the statements below best describes the strategy of Apriori in finding the frequent itemsets?

**Scegli un'alternativa:**

- ☒ a. Evaluation of the support of the itemsets in an order such that uninteresting parts of the search space are considered only at the end of the execution
- ☒ b. Evaluation of the confidence of the itemsets in an order such that uninteresting parts of the search space are pruned as soon as possible
- ☒ c. Evaluation of the support of the itemsets in an order such that uninteresting parts of the search space are pruned as soon as possible
- ☒ d. Evaluation of the support of the itemsets in an order such that the interesting parts of the search space are pruned as soon as possible

Domanda **12**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Consider the transactional dataset below

**ID Items**

- 1 A,B,C
- 2 A,B,D
- 3 B,D,E
- 4 C,D
- 5 A,C,D,E

Which is the *support* of the rule  $A, C \Rightarrow B$ ?

**Scegli un'alternativa:**

- ☒ a. 100%
- ☐ b. 40%
- ☐ c. 20%
- ☐ d. 50%

Domanda **13**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

In order to reduce the dimensionality of a dataset, which is the advantage of Multi Dimensional Scaling (MDS), with respect to Principal Component Analysis (PCA)

**Scegli un'alternativa:**

- ☒ a. MDS can be used also with categorical data, provided that the matrix of the distance is available, while PCA is limited to vector spaces
- ☐ b. MDS requires less computational power
- ☐ c. MDS can be used with categorical data after a transformation in a vector space
- ☐ d. MDS can work on any kind of data, while PCA is limited to categorical data

Domanda **14**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

When is polynomial regression appropriate?

- ☐ a. When there is more than one predicting attribute
- ☐ b. When the relationship between the predicting variable and the target cannot be approximated as linear
- ☐ c. When it is necessary to project the data into a higher dimensional space
- ☒ d. When the target values are not linearly separable



Domanda **15**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Which is different from the others?

**Scegli un'alternativa:**

- ☒ a. Decision Tree
- ☐ b. K-means
- ☐ c. Expectation Maximisation
- ☐ d. Apriori

Vai a...

[Machine Learning - Python Lab](#) ►